Midrange Computing Workshop

Sandy Merola

Gary Jung

March 26, 2002

Approach

- Survey results

- Options:
  - Support
  - Shared computational resources

- Open discussion

- Firm next steps

Survey: Received 43 Responses

Environmental Energy Technologies	7
AFRD	7
Nuclear Science	5
Physics	5
NERSC	4
Physical Biosciences	4
Chemical Sciences	3
Life Sciences	3
Material Sciences	3
Earth Sciences	2

No response from ALS and Genome

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Type of Research

Experimental	12
NS, HEP	
Simulation/Modeling	12
EETD, AFRD, ESD, PBD, LSD	
Theory	9
CSD, AFRD, MS, EETD	

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Current Primary Computing System

Linux, Mac, SGI, Solaris, Compaq Alpha desktops	26
PDSF Physics, NS	9
NERSC IBM SP Utilization	4
Linux Clusters	2
18 Processor IBM Power 3 Cluster	1
Cray T3E	1

3/26/02

Impact of
Increased Computing Resources

Analyze larger volume of data	16
Analyze experimental data faster	19
Perform larger simulations	20
Perform faster simulations	27
Perform simulations with higher resolutions	19
Implement new alogrithms resulting in improved simulations	18

Almost all Physics, NS, PBD would use high performance computing to do larger volumes and analyze data faster

Form of Computing
That Would Be Most Useful

Medium Cluster	16
Medium size SMP	15
High Performance Desktop	4
Other	1

3/26/02

Critical Elements In A New System

Memory size	25
Processor clockspeed	25
Storage	20
Network connectivity	16
I/O	14
Tightly coupled processors	9

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Source of Software

Written by group	27
Freely available	8
Commercial	6

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Midrange Computing Readiness

Ready now	17
Will be ready shortly	7
Will be ready mid-term	7
Will be ready long-term	3
Unsure	8

3/26/02

How Parallelizable Is Your Code?

Already done	12
Easy	5
Moderately difficult	5
Difficult	6
Inconceivable	1
Unnecessary, serial OK	11
Unsure	3

Memory Model – most respondents indicated either distributed or shared could be accommodated; many didn't know

LAWRENCE BERKELEY NATIONAL LABORATORY

Planned Procurements

Linux cluster	13
Expansion of current clusters	2
SMP consideration	2
No change	3
Unsure	23

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Support

Prepurchase consulting	17
Vendor negotiating expertise	13
Facilities	20
Configuration expertise	25
HW maintenance	22
Ongoing support	25
Application porting support	8

Comments

- Quality of support

- Cost of support (reasonable)

- Leveraging NERSC

- Networking infrastructure

- In the case of a pooled or institutional usage, it is important to determine the appropriate size of the shared resource

So, now we discuss support options

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Pertinent Issues for Support

Standardization

- Cannot fully realize economies of scale if clusters are different
- More difficult to manage a cluster built by someone else

Scale

- ITSD currently supports 2 small clusters and is willing to develop a service offering
- Support of larger clusters would require the Laboratory to develop the expertise

3/26/02

Pre-Purchase Consulting

Deliver the basics for RFP

What can we provide?

- Advice on small to mid size clusters up to 32 nodes
  (more complex at > 32 nodes – e.g., network switch latency issues)
- ITSD might setup a small cluster to provide a
  "try before you buy" service
- Cost analysis of purchase, timeline, and effort
- Specifying systems HW configuration or components
- Specifying peripherals such as racks, UPS, kvm terminal switches
- Specifying cluster distribution
- Estimating software licensing costs
- Recommendations for data storage systems
- Vendor recommendations

Computer Room Space

What is the advantage of a centralized Facility?

Machine room environment:

- Access to electrical infrastructure
- Proper air conditioning
- Access to high speed local area & wide area networks
- Secure card key access

3/26/02

Facilities: Examples of Costs

One time costs:

Transportation, seismic bracing, electrical	$1,000
LBNL Network Drop	$400 per drop
Facilities Coordination (1.5 – 2 days)	$1,500

Recurring Costs:

Housing costs in either 50A-2109 or 50B-1275 Computer Room per rack (including space & electricity)	$225/rack/mo

Initial Set Up and Configuration

Major set up tasks:

- Assembly of racks and equipment
- HW assembly and network wiring
- Build master node, set up file systems
- Install PGI compilers
- Integration of 3rd party compilers (Portland Group)
- Build Myrinet drivers/kernel modules
- Build client image
- Install client node file systems

Example: Estimate of effort for a 10 node system with Myrinet, PGI compilers: 3 days

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Hardware Maintenance

- PC hardware tends to be less reliable, especially on larger clusters

- Important to get a responsible vendor

- Users with larger clusters should consider purchasing spares

3/26/02

Systems & Security Administration

What does CIS provide:

- Upgrades
- Updating of nodes
- Security/SSH
- Troubleshooting
- Crash recovery
- User account admin
- Network admin – sendmail, NFS
- Installation of 3rd party software
- Software license management
- Scheduler
- Monitoring of nodes

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Advantages of Institutional Set Up and Support

- Better coverage, expertise
- Expertise, knowledge
- Economy of scale
- Best practice
- Standardization
- Can mean days instead of weeks for troubleshooting
- Cyber protection and emergency response

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Cost Factors

What are the cost factors in providing ongoing systems admin?

- # of cluster nodes
- # of users
- Is the system used for code development or production running?

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Effort

What is the level of effort to provide system admin support?

	Minimal Level	Standard Level
10 node cluster w/ 1 master node	1.5 days/mo	3 days/mo
11-20 node cluster w/ 1 master node	2 days	4 days
21-30 node cluster w/ 1 master node	2.5 days/mo	5 days/mo

* Current effort costs are $110/hr or $880/day

Feasibility

Some issues may not be feasible for us to address (outside our core competency at this time)

- Determining if code is suitable to run on a cluster
- Defining classes of problems – some may run better depending on cluster configuration
- Porting issues: How do we marry code to cluster?
- Formal procurement/negotiations

3/26/02

Shared Computational Resource

- 20 respondents indicated they may be interested in pooling resources with another project to gain access to a larger system or lower support costs

- Same respondents would also be interested in pooling with several projects

- Approximately 15 of 17 respondents who are considering procurement, stated a preference for a cluster

3/26/02

Shared Resource Options

1. No offering at this time
   - Acceptable
2. Provide systems support as a gradual mechanism to create a shared resource
3. Procure an institutional MRC
4. Build on an existing computational resource
   - alvarez, PDSF, or division owned

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Shared Resource

A shared mid-range computing resource must be:

- Appropriate
- Sustainable

This implies:

- Compatible user requirements
- Advantage to the programs
- Affordable acquisition
- Sustainable financial model

Issues

There must be an added-value that results from sharing before divisions/projects would be willing to give up control of owning/running their own systems

- Cheaper
- Expertise
- Environment
- Fungibility of resources
- Cybersecurity

If ITSD were to facilitate this, it must build expertise to provide added-value

- $
- Time

3/26/02

Issues

- Under any approach, there is an institutional startup cost for shared resource

- A combined and shared resource could be managed to provide a more powerful resource than the same capability owned and controlled individually

- Bky Lab management must see an institutional advantage in order to allocate overhead dollars

3/26/02

Growing A Shared Resource

- Systems support may be a gradual means of creating an shared resource

- Fungible resource could allow building/sharing of a larger machine given future divisional investments

- Lab overhead might help with this, if a large institutional advantage can be recognized

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02

Procure an Institutional MRC

- A number of divisions could contribute to the acquisition and startup costs of a new MRC

Build On Existing Computational Resources

Discussion:

- What could be the role of PDSF?

- What could be the role of alvarez?

- Is there an existing divisional owned computer that could serve as the foundation for growing a shared resource?

- Other pertinent questions?

LAWRENCE BERKELEY NATIONAL LABORATORY

Path Forward

- ITSD will provide a specific acquisition and/or support proposal at your invitation

- If there is sufficient interest, ITSD will facilitate a working group that will result in the creation of a shared resource

LAWRENCE BERKELEY NATIONAL LABORATORY

3/26/02